Knowledge-Based Scene Graph Generation in Medical Field

Jessica D'souza
Department of CSE
St. Joseph Engineering College
Vamanjoor, Karnataka, India
dsouzjess17@gmail.com

Aleema P. K.
Department of CSE
St. Joseph Engineering College
Vamanjoor, Karnataka, India
aleemaparakatta@gmail.com

Dhanyashree S.
Department of CSE
St. Joseph Engineering College
Vamanjoor, Karnataka, India
dhanyashrees717@gmail.com

Clita Fernandes
Department of CSE
St. Joseph Engineering College
Vamanjoor, Karnataka, India
clitafernandes01@gmail.com

K. M. Kavitha
Department of CSE
St. Joseph Engineering College
Vamanjoor, Karnataka, India
kavithakmahesh@gmail.com

Chandra Naik
Department of CSE
St. Joseph Engineering College
Vamanjoor, Karnataka, India
chandra.nitk2017@gmail.com

***Abstract*— This work aims to combine object detection and knowledge graphs to understand medical scenes. The rapid growth of the medical field has introduced various equipment types which makes scene analysis more crucial. By representing object detection results as a knowledge graph, one can derive conclusions and gain a clear understanding of the scene through scene graphs. While object detection identifies objects in images, it lacks the ability to utilize contextual information and visual reason. Therefore, knowledge-aware object detection becomes essential in allowing the integration of external knowledge graphs into object detection algorithms. Therefore, the work focuses on the medical field that combines object detection and object relationships to form a correlated understanding of the scene. Construct relationships between objects, generate scene graphs, and derive inferences to predict the overall scene understanding is achieved through the knowledge graph. The proposed work incorporates knowledge graphs in the medical domain that addresses the need for improved scene understanding in medical contexts.**

Keywords—Scene Graph, Medical Field, Object Detection, Faster R-CNN, YOLO Architecture

I. INTRODUCTION

Scene understanding in the medical field is based on visual context using object detection and representing it in the form of knowledge graphs in order to derive conclusions or understand contexts.

Object detection has been very common with the advent of artificial intelligence and machine learning. Detecting objects from image data with a bounding box is now made possible by several machine learning algorithms, neural network algorithms, and with the increasing popularity of deep learning models. But classical object detection frameworks lack the ability to utilize objects surrounding information. The ability to understand the context from the given input scenario and to be able to visually reason is a major part of ongoing research.

In the proposed work, we restrict this application to the medical field where there is a broader opportunity for scene understanding. Visual and contextual semantic features can both assist models in accurately predicting object relationships and are combined to best learn the global contextual information contained in a scene. We implement this by considering the statistical correlations between object pairs as language priors. Thus, with these scene graphs, we can extract contextual information after incorporating auxiliary information based on external knowledge graphs and predict the overall understanding of the scene/ input image data.

Scene graphs can be used in medical robots as well as surgical planning as they allow surgeons to perform surgery with great accuracy and precision. This also enables minimalistic planning. It can also be used in medical diagnosis since it can provide a detailed representation of a patient's anatomy. Moreover, it may help the visually impaired to understand what is in front of them. Scene graphs help in medical education wherein the students can learn the relationship between different anatomical structures as well as visualize complex medical concepts.

II. LITERATURE SURVEY

The advancements in computer vision and artificial intelligence, numerous works have been done in the knowledge graphs. Fang et al. introduce a framework that incorporates external knowledge from knowledge graphs into object detection algorithms. The authors present two approaches, frequency-based and knowledge graph-based to quantify background knowledge for improved detections [1]. While the frequency-based method holds promise, the effectiveness of the knowledge graph-based approach is not reproducible across different object detection models. The paper emphasizes the importance of tailored implementation for specific object detection algorithms and highlights the benefits of this approach for lower-scoring detections with multiple detections per image. However, the practical applicability of the framework to state-of-the-art object detection challenges is limited. The authors suggest exploring the potential of semantic consistency from background knowledge in other research areas, such as action anticipation and scene understanding. Although the paper lacks code implementation and sufficient details on the dataset, additional research and analysis were conducted to complement and address the gaps in the paper's findings.

Wang et al. in [2] discuss the implementation of knowledge graphs in object detection. They emphasize the generation of knowledge graphs for scene understanding and propose a framework that combines visual information and knowledge reasoning methods using Faster R-CNN and KG-based reasoning. They focus on few-shot object detection, where limited samples of new objects are used with prior knowledge from base classes. The paper describes the extraction of a knowledge subgraph from a large knowledge graph to infer unknown objects and enhance semantic

2023 IEEE International Conference on Distributed Computing, VLSI, Electrical Circuits and Robotics (DISCOVER) | 979-8-3503-4012-9/23/$31.00 ©2023 IEEE | DOI: 10.1109/DISCOVER58630.2023.10316713

979-8-3503-4012-9/23/$31.00 ©2023 IEEE